Chapter 3

Determinants

3.1 Permutations, Signature of a Permutation

This chapter contains a review of determinants and their use in linear algebra. We begin with permutations and the signature of a permutation. Next, we define multilinear maps and alternating multilinear maps. Determinants are introduced as alternating multilinear maps taking the value 1 on the unit matrix (following Emil Artin). It is then shown how to compute a determinant using the Laplace expansion formula, and the connection with the usual definition is made. It is shown how determinants can be used to invert matrices and to solve (at least in theory!) systems of linear equations (the Cramer formulae). Finally, the determinant of a linear map is defined.

Determinants can be defined in several ways. For example, determinants can be defined in a fancy way in terms of the exterior algebra (or alternating algebra) of a vector space. We will follow a more algorithmic approach due to Emil Artin. No matter which approach is followed, we need a few preliminaries about permutations on a finite set. We need to show that every permutation on n elements is a product of transpositions, and that the parity of the number of transpositions involved is an invariant of the permutation. Let $I_n = \{1, 2 \ldots, n\}$, where $n \in \mathbb{N}$, and $n > 0$.

Definition 3.1 A *permutation on n elements* is a bijection $\pi \colon I_n \to I_n$. When $n = 1$, the only function from $\{1\}$ to $\{1\}$ is the constant map $1 \mapsto 1$. Thus, we will assume that $n \geq 2$. A *transposition* is a permutation $\tau \colon I_n \to I_n$ such that, for some $i < j$ (with $1 \leq i < j \leq n$), $\tau(i) = j$, $\tau(j) = i$, and $\tau(k) = k$, for all $k \in I_n - \{i, j\}$. In other words, a transposition exchanges two distinct elements $i, j \in I_n$. A *cyclic permutation of order k (or k-cycle)* is a permutation $\sigma \colon I_n \to I_n$ such that, for some $i_1, i_2, \ldots, i_k$, with $1 \leq i_1 < i_2 < \ldots < i_k \leq n$, and $k \geq 2$,

$$\sigma(i_1) = i_2, \ldots, \sigma(i_{k-1}) = i_k,\ \sigma(i_k) = i_1,$$

and $\sigma(j) = j$, for $j \in I_n - \{i_1, \ldots, i_k\}$. The set $\{i_1, \ldots, i_k\}$ is called the *domain* of the cyclic permutation, and the cyclic permutation is sometimes denoted by $(i_1, i_2, \ldots, i_k)$.

If τ is a transposition, clearly, $\tau \circ \tau = \mathrm{id}$. Also, a cyclic permutation of order 2 is a transposition, and for a cyclic permutation σ of order k, we have $\sigma^k = \mathrm{id}$. The following proposition shows the importance of cyclic permutations and transpositions. We will also use the terminology product of permutations (or transpositions), as a synonym for composition of permutations.

Proposition 3.1 *For every $n \geq 2$, for every permutation $\pi\colon I_n \to I_n$, there is a partition of I_n into r subsets, with $1 \leq r \leq n$, where each set J in this partition is either a singleton $\{i\}$, or it is of the form*

$$J = \{i, \pi(i), \pi^2(i), \ldots, \pi^{r_i-1}(i)\},$$

where r_i is the smallest integer, such that, $\pi^{r_i}(i) = i$ and $2 \leq r_i \leq n$. If π is not the identity, then it can be written in a unique way as a composition $\pi = \sigma_1 \circ \ldots \circ \sigma_s$ of cyclic permutations (where $1 \leq s \leq r$). Every permutation $\pi\colon I_n \to I_n$ can be written as a nonempty composition of transpositions.

Proof. Consider the relation R_π defined on I_n as follows: $iR_\pi j$ iff there is some $k \geq 1$ such that $j = \pi^k(i)$. We claim that R_π is an equivalence relation. Transitivity is obvious. We claim that for every $i \in I_n$, there is some least r ($1 \leq r \leq n$) such that $\pi^r(i) = i$. Indeed, consider the following sequence of $n+1$ elements:

$$\langle i, \pi(i), \pi^2(i), \ldots, \pi^n(i)\rangle.$$

Since I_n only has n distinct elements, there are some h, k with $0 \leq h < k \leq n$ such that

$$\pi^h(i) = \pi^k(i),$$

and since π is a bijection, this implies $\pi^{k-h}(i) = i$, where $0 \leq k - h \leq n$. Thus, R_π is reflexive. It is symmetric, since if $j = \pi^k(i)$, letting r be the least $r \geq 1$ such that $\pi^r(i) = i$, then

$$i = \pi^{kr}(i) = \pi^{k(r-1)}(\pi^k(i)) = \pi^{k(r-1)}(j).$$

Now, for every $i \in I_n$, the equivalence class of i is a subset of I_n, either the singleton $\{i\}$ or a set of the form

$$J = \{i, \pi(i), \pi^2(i), \ldots, \pi^{r_i-1}(i)\},$$

where r_i is the smallest integer such that $\pi^{r_i}(i) = i$ and $2 \leq r_i \leq n$, and in the second case, the restriction of π to J induces a cyclic permutation σ_i, and $\pi = \sigma_1 \circ \ldots \circ \sigma_s$, where s is the number of equivalence classes having at least two elements.

For the second part of the proposition, we proceed by induction on n. If $n = 2$, there are exactly two permutations on $\{1, 2\}$, the transposition τ exchanging 1 and 2, and the identity. However, $\mathrm{id}_2 = \tau^2$. Now, let $n \geq 3$. If $\pi(n) = n$, since by the induction hypothesis, the restriction of π to $\{1, \ldots, n-1\}$ can be written as a product of transpositions, π itself can be written as a product of transpositions. If $\pi(n) = k \neq n$, letting τ be the transposition such that $\tau(n) = k$ and $\tau(k) = n$, it is clear that $\tau \circ \pi$ leaves n invariant, and by the induction hypothesis, we have $\tau \circ \pi = \tau_m \circ \ldots \circ \tau_1$ for some transpositions, and thus

$$\pi = \tau \circ \tau_m \circ \ldots \circ \tau_1,$$

a product of transpositions (since $\tau \circ \tau = \mathrm{id}_n$). $\square$

Remark: When $\pi = \mathrm{id}_n$ is the identity permutation, we can agree that the composition of 0 transpositions is the identity. The second part of Proposition 3.1 shows that the transpositions generate the set of all permutations.

In writing a permutation π as a composition $\pi = \sigma_1 \circ \ldots \circ \sigma_s$ of cyclic permutations, it is clear that the order of the σ_i does not matter, since their domains are disjoint. Given a permutation written as a product of transpositions, we now show that the parity of the number of transpositions is an invariant.

Definition 3.2 For every $n \geq 2$, since every permutation $\pi\colon I_n \to I_n$ defines a partition of r subsets over which π acts either as the identity or as a cyclic permutation, let $\epsilon(\pi)$, called the *signature* of π, be defined by $\epsilon(\pi) = (-1)^{n-r}$, where r is the number of sets in the partition.

If τ is a transposition exchanging i and j, it is clear that the partition associated with τ consists of $n-1$ equivalence classes, the set $\{i, j\}$, and the $n-2$ singleton sets $\{k\}$, for $k \in I_n - \{i, j\}$, and thus, $\epsilon(\tau) = (-1)^{n-(n-1)} = (-1)^1 = -1$.

Proposition 3.2 *For every $n \geq 2$, for every permutation $\pi\colon I_n \to I_n$, for every transposition τ, we have*

$$\epsilon(\tau \circ \pi) = -\epsilon(\pi).$$

Consequently, for every product of transpositions such that $\pi = \tau_m \circ \ldots \circ \tau_1$, we have

$$\epsilon(\pi) = (-1)^m,$$

which shows that the parity of the number of transpositions is an invariant.

Proof. Assume that $\tau(i) = j$ and $\tau(j) = i$, where $i < j$. There are two cases, depending whether i and j are in the same equivalence class J_l of R_π, or if they are in distinct equivalence classes. If i and j are in the same class J_l, then if

$$J_l = \{i_1, \ldots, i_p, \ldots i_q, \ldots i_k\},$$

where $i_p = i$ and $i_q = j$, since

$$\tau(\pi(\pi^{-1}(i_p))) = \tau(i_p) = \tau(i) = j = i_q$$

and

$$\tau(\pi(i_{q-1})) = \tau(i_q) = \tau(j) = i = i_p,$$

it is clear that J_l splits into two subsets, one of which is $\{i_p, \ldots, i_{q-1}\}$, and thus, the number of classes associated with $\tau \circ \pi$ is $r + 1$, and $\epsilon(\tau \circ \pi) = (-1)^{n-r-1} = -(-1)^{n-r} = -\epsilon(\pi)$. If i and j are in distinct equivalence classes J_l and J_m, say

$$\{i_1, \ldots, i_p, \ldots i_h\}$$

and

$$\{j_1, \ldots, j_q, \ldots j_k\},$$

where $i_p = i$ and $j_q = j$, since

$$\tau(\pi(\pi^{-1}(i_p))) = \tau(i_p) = \tau(i) = j = j_q$$

and

$$\tau(\pi(\pi^{-1}(j_q))) = \tau(j_q) = \tau(j) = i = i_p,$$

we see that the classes J_l and J_m merge into a single class, and thus, the number of classes associated with $\tau \circ \pi$ is $r - 1$, and $\epsilon(\tau \circ \pi) = (-1)^{n-r+1} = -(-1)^{n-r} = -\epsilon(\pi)$.

Now, let $\pi = \tau_m \circ \ldots \circ \tau_1$ be any product of transpositions. By the first part of the proposition, we have

$$\epsilon(\pi) = (-1)^{m-1}\epsilon(\tau_1) = (-1)^{m-1}(-1) = (-1)^m,$$

since $\epsilon(\tau_1) = -1$ for a transposition. $\square$

Remark: When $\pi = \mathrm{id}_n$ is the identity permutation, since we agreed that the composition of 0 transpositions is the identity, it it still correct that $(-1)^0 = \epsilon(\mathrm{id}) = +1$. From the proposition, it is immediate that $\epsilon(\pi' \circ \pi) = \epsilon(\pi')\epsilon(\pi)$. In particular, since $\pi^{-1} \circ \pi = \mathrm{id}_n$, we get $\epsilon(\pi^{-1}) = \epsilon(\pi)$.

We can now proceed with the definition of determinants.

3.2 Alternating Multilinear Maps

First, we define multilinear maps, symmetric multilinear maps, and alternate multilinear maps.

Remark: Most of the definitions and results presented in this section also hold when K is a commutative ring, and when we consider modules over K (free modules, when bases are needed).

Let $E_1, \ldots, E_n$, and F, be vector spaces over a field K, where $n \geq 1$.

Definition 3.3 A function $f\colon E_1 \times \ldots \times E_n \to F$ is a *multilinear map (or an n-linear map)* if it is linear in each argument, holding the others fixed. More explicitly, for every i, $1 \leq i \leq n$, for all $x_1 \in E_1 \ldots$, $x_{i-1} \in E_{i-1}$, $x_{i+1} \in E_{i+1}, \ldots$, $x_n \in E_n$, for all $x, y \in E_i$, for all $\lambda \in K$,

$$
\begin{aligned}
&f(x_1, \ldots, x_{i-1}, x+y, x_{i+1}, \ldots, x_n) = \\
&\qquad\qquad f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n), \\
&f(x_1, \ldots, x_{i-1}, \lambda x, x_{i+1}, \ldots, x_n) = \lambda f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n).
\end{aligned}
$$

When $F = K$, we call f an *n-linear form (or multilinear form)*. If $n \geq 2$ and $E_1 = E_2 = \ldots = E_n$, an n-linear map $f\colon E \times \ldots \times E \to F$ is called *symmetric*, if $f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)})$, for every permutation π on $\{1, \ldots, n\}$. An n-linear map $f\colon E \times \ldots \times E \to F$ is called *alternating*, if $f(x_1, \ldots, x_n) = 0$ whenever $x_i = x_{i+1}$, for some i, $1 \leq i \leq n-1$ (in other words, when two adjacent arguments are equal). It does not harm to agree that when $n = 1$, a linear map is considered to be both symmetric and alternating, and we will do so.

When $n = 2$, a 2-linear map $f\colon E_1 \times E_2 \to F$ is called a *bilinear map*. We have already seen several examples of bilinear maps. Multiplication $\cdot\colon K \times K \to K$ is a bilinear map, treating K as a vector space over itself. More generally, multiplication $\cdot\colon A \times A \to A$ in a ring A is a bilinear map, viewing A as a module over itself.

The operation $\langle -, - \rangle\colon E^* \times E \to K$ applying a linear form to a vector is a bilinear map.

Symmetric bilinear maps (and multilinear maps) play an important role in geometry (inner products, quadratic forms), and in differential calculus (partial derivatives).

A bilinear map is symmetric if $f(u, v) = f(v, u)$, for all $u, v \in E$.

Alternating multilinear maps satisfy the following simple but crucial properties.

Proposition 3.3 *Let $f\colon E \times \ldots \times E \to F$ be an n-linear alternating map, with $n \geq 2$. The following properties hold:*

(1)
$$f(\ldots, x_i, x_{i+1}, \ldots) = -f(\ldots, x_{i+1}, x_i, \ldots)$$

(2)
$$f(\ldots, x_i, \ldots, x_j, \ldots) = 0,$$
where $x_i = x_j$, and $1 \leq i < j \leq n$.

(3)
$$f(\ldots, x_i, \ldots, x_j, \ldots) = -f(\ldots, x_j, \ldots, x_i, \ldots),$$
where $1 \leq i < j \leq n$.

(4)

$$f(\ldots, x_i, \ldots) = f(\ldots, x_i + \lambda x_j, \ldots),$$

for any $\lambda \in K$, and where $i \neq j$.

Proof. (1) By multilinearity applied twice, we have

$$\begin{aligned} f(\ldots, x_i + x_{i+1}, x_i + x_{i+1}, \ldots) = f(\ldots, x_i, x_i, \ldots) + f(\ldots, x_i, x_{i+1}, \ldots) \\ + f(\ldots, x_{i+1}, x_i, \ldots) + f(\ldots, x_{i+1}, x_{i+1}, \ldots), \end{aligned}$$

and since f is alternating, this yields

$$0 = f(\ldots, x_i, x_{i+1}, \ldots) + f(\ldots, x_{i+1}, x_i, \ldots),$$

that is, $f(\ldots, x_i, x_{i+1}, \ldots) = -f(\ldots, x_{i+1}, x_i, \ldots)$.

(2) If $x_i = x_j$ and i and j are not adjacent, we can interchange x_i and x_{i+1}, and then x_i and x_{i+2}, etc, until x_i and x_j become adjacent. By (1),

$$f(\ldots, x_i, \ldots, x_j, \ldots) = \epsilon f(\ldots, x_i, x_j, \ldots),$$

where $\epsilon = +1$ or -1, but $f(\ldots, x_i, x_j, \ldots) = 0$, since $x_i = x_j$, and (2) holds.

(3) follows from (2) as in (1). (4) is an immediate consequence of (2). $\square$

Proposition 3.3 will now be used to show a fundamental property of alternating linear maps. First, we need to extend the matrix notation a little bit. Let E be a vector space over K. Given an $n \times n$ matrix $A = (a_{i\,j})$ over K, we can define a map $L(A)\colon E^n \to E^n$ as follows:

$$\begin{aligned} L(A)_1(u_1) &= a_{1\,1} u_1 + \cdots + a_{1\,n} u_n, \\ &\ldots \\ L(A)_n(u_n) &= a_{n\,1} u_1 + \cdots + a_{n\,n} u_n, \end{aligned}$$

for all $u_1, \ldots, u_n \in E$. It is immediately verified that $L(A)$ is linear. Then, given two $n \times n$ matrice $A = (a_{i\,j})$ and $B = (b_{i\,j})$, by repeating the calculations establishing the product of matrices (just before Definition 2.12), we can show that

$$L(AB) = L(A) \circ L(B).$$

It is then convenient to use the matrix notation to describe the effect of the linear map $L(A)$, as

$$\begin{pmatrix} L(A)_1(u_1) \\ L(A)_2(u_2) \\ \vdots \\ L(A)_n(u_n) \end{pmatrix} = \begin{pmatrix} a_{1\,1} & a_{1\,2} & \ldots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \ldots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n\,1} & a_{n\,2} & \ldots & a_{n\,n} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Lemma 3.4 *Let $f\colon E \times \ldots \times E \to F$ be an n-linear alternating map. Let $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ be two families of n vectors, such that,*

$$\begin{aligned} v_1 &= a_{1\,1}u_1 + \cdots + a_{n\,1}u_n, \\ &\ldots \\ v_n &= a_{1\,n}u_1 + \cdots + a_{n\,n}u_n. \end{aligned}$$

Equivalently, letting

$$A = \begin{pmatrix} a_{1\,1} & a_{1\,2} & \ldots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \ldots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n\,1} & a_{n\,2} & \ldots & a_{n\,n} \end{pmatrix}$$

assume that we have

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = A^\top \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Then,

$$f(v_1, \ldots, v_n) = \Big(\sum_\pi \epsilon(\pi) a_{\pi(1)\,1} \cdots a_{\pi(n)\,n}\Big) f(u_1, \ldots, u_n),$$

where the sum ranges over all permutations π on $\{1, \ldots, n\}$.

Proof. Expanding $f(v_1, \ldots, v_n)$ by multilinearity, we get a sum of terms of the form

$$a_{\pi(1)\,1} \cdots a_{\pi(n)\,n} f(u_{\pi(1)}, \ldots, u_{\pi(n)}),$$

for all possible functions $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$. However, because f is alternating, only the terms for which π is a permutation are nonzero. By Proposition 3.1, every permutation π is a product of transpositions, and by Proposition 3.2, the parity $\epsilon(\pi)$ of the number of transpositions only depends on π. Then, applying Proposition 3.3 (3) to each transposition in π, we get

$$a_{\pi(1)\,1} \cdots a_{\pi(n)\,n} f(u_{\pi(1)}, \ldots, u_{\pi(n)}) = \epsilon(\pi) a_{\pi(1)\,1} \cdots a_{\pi(n)\,n} f(u_1, \ldots, u_n).$$

Thus, we get the expression of the lemma. □

The quantity

$$\det(A) = \sum_\pi \epsilon(\pi) a_{\pi(1)\,1} \cdots a_{\pi(n)\,n}$$

is in fact the value of the determinant of A (which, as we shall see shortly, is also equal to the determinant of $A^\top$). However, working directly with the above definition is quite ackward, and we will proceed via a slightly indirect route

3.3 Definition of a Determinant

Recall that the set of all square $n \times n$-matrices with coefficients in a field K is denoted by $\mathrm{M}_n(K)$.

Definition 3.4 A determinant is defined as any map

$$D\colon \mathrm{M}_n(K) \to K,$$

which, when viewed as a map on $(K^n)^n$, i.e., a map of the n columns of a matrix, is n-linear alternating and such that $D(I_n) = 1$ for the identity matrix I_n. Equivalently, we can consider a vector space E of dimension n, some fixed basis $(e_1, \ldots, e_n)$, and define

$$D\colon E^n \to K$$

as an n-linear alternating map such that $D(e_1, \ldots, e_n) = 1$.

First, we will show that such maps D exist, using an inductive definition that also gives a recursive method for computing determinants. Actually, we will define a family $(\mathcal{D}_n)_{n\geq 1}$ of (finite) sets of maps $D\colon \mathrm{M}_n(K) \to K$. Second, we will show that determinants are in fact uniquely defined, that is, we will show that each $\mathcal{D}_n$ consists of a single map. This will show the equivalence of the direct definition $\det(A)$ of Lemma 3.4 with the inductive definition $D(A)$. Finally, we will prove some basic properties of determinants, using the uniqueness theorem. Given a matrix $A \in \mathrm{M}_n(K)$, we denote its n columns by $A^1, \ldots, A^n$.

Definition 3.5 For every $n \geq 1$, we define a finite set $\mathcal{D}_n$ of maps $D\colon \mathrm{M}_n(K) \to K$ inductively as follows:

When $n = 1$, $\mathcal{D}_1$ consists of the single map D such that, $D(A) = a$, where $A = (a)$, with $a \in K$.

Assume that $\mathcal{D}_{n-1}$ has been defined, where $n \geq 2$. We define the set $\mathcal{D}_n$ as follows. For every matrix $A \in \mathrm{M}_n(K)$, let $A_{i\,j}$ be the $(n-1) \times (n-1)$-matrix obtained from $A = (a_{i\,j})$ by deleting row i and column j. Then, $\mathcal{D}_n$ consists of all the maps D such that, for some i, $1 \leq i \leq n$,

$$D(A) = (-1)^{i+1} a_{i\,1} D(A_{i\,1}) + \cdots + (-1)^{i+n} a_{i\,n} D(A_{i\,n}),$$

where for every j, $1 \leq j \leq n$, $D(A_{i\,j})$ is the result of applying any D in $\mathcal{D}_{n-1}$ to $A_{i\,j}$.

We confess that the use of the same letter D for the member of $\mathcal{D}_n$ being defined, and for members of $\mathcal{D}_{n-1}$, may be slightly confusing. We considered using subscripts to distinguish, but this seems to complicate things unnecessarily. One should not worry too much anyway, since it will turn out that each $\mathcal{D}_n$ contains just one map.

Each $(-1)^{i+j} D(A_{i\,j})$ is called the *cofactor* of $a_{i\,j}$, and the inductive expression for $D(A)$ is called a *Laplace expansion of* D *according to the* i*-th row*. Given a matrix $A \in \mathrm{M}_n(K)$, each $D(A)$ is called a *determinant* of A.

We will prove shortly that $D(A)$ is uniquely defined (at the momement, it is not clear that $\mathcal{D}_n$ consists of a single map). Assuming this fact, given a $n \times n$-matrix $A = (a_{i\,j})$,

$$A = \begin{pmatrix} a_{1\,1} & a_{1\,2} & \dots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \dots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n\,1} & a_{n\,2} & \dots & a_{n\,n} \end{pmatrix}$$

its determinant is denoted by $D(A)$ or $\det(A)$, or more explicitly by

$$\det(A) = \begin{vmatrix} a_{1\,1} & a_{1\,2} & \dots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \dots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n\,1} & a_{n\,2} & \dots & a_{n\,n} \end{vmatrix}$$

First, let us first consider some examples.

Example 3.1

1. When $n = 2$, if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

expanding according to any row, we have

$$D(A) = ad - bc.$$

2. When $n = 3$, if

$$A = \begin{pmatrix} a_{1\,1} & a_{1\,2} & a_{1\,3} \\ a_{2\,1} & a_{2\,2} & a_{2\,3} \\ a_{3\,1} & a_{3\,2} & a_{3\,3} \end{pmatrix}$$

expanding according to the first row, we have

$$D(A) = a_{1\,1} \begin{vmatrix} a_{2\,2} & a_{2\,3} \\ a_{3\,2} & a_{3\,3} \end{vmatrix} - a_{1\,2} \begin{vmatrix} a_{2\,1} & a_{2\,3} \\ a_{3\,1} & a_{3\,3} \end{vmatrix} + a_{1\,3} \begin{vmatrix} a_{2\,1} & a_{2\,2} \\ a_{3\,1} & a_{3\,2} \end{vmatrix}$$

that is,

$$D(A) = a_{1\,1}(a_{2\,2}a_{3\,3} - a_{3\,2}a_{2\,3}) - a_{1\,2}(a_{2\,1}a_{3\,3} - a_{3\,1}a_{2\,3}) + a_{1\,3}(a_{2\,1}a_{3\,2} - a_{3\,1}a_{2\,2}),$$

which gives the explicit formula

$$D(A) = a_{1\,1}a_{2\,2}a_{3\,3} + a_{2\,1}a_{3\,2}a_{1\,3} + a_{3\,1}a_{1\,2}a_{2\,3} - a_{1\,1}a_{3\,2}a_{2\,3} - a_{2\,1}a_{1\,2}a_{3\,3} - a_{3\,1}a_{2\,2}a_{1\,3}.$$

We now show that each $D \in \mathcal{D}_n$ is a determinant (map).

Lemma 3.5 *For every $n \geq 1$, for every $D \in \mathcal{D}_n$ as defined in Definition 3.5, D is an alternate linear map such that $D(I_n) = 1$.*

Proof. By induction on n, it is obvious that $D(I_n) = 1$. Let us now prove that D is multilinear. Let us show that D is linear in each column. Consider any column k. Since

$$D(A) = (-1)^{i+1} a_{i\,1} D(A_{i\,1}) + \cdots + (-1)^{i+j} a_{i\,j} D(A_{i\,j}) + \cdots + (-1)^{i+n} a_{i\,n} D(A_{i\,n}),$$

if $j \neq k$, then by induction, $D(A_{i\,j})$ is linear in column k, and $a_{i\,j}$ does not belong to column k, so $(-1)^{i+j} a_{i\,j} D(A_{i\,j})$ is linear in column k. If $j = k$, then $D(A_{i\,j})$ does not depend on column $k = j$, since $A_{i\,j}$ is obtained from A by deleting row i and column $j = k$, and $a_{i\,j}$ belongs to column $j = k$. Thus, $(-1)^{i+j} a_{i\,j} D(A_{i\,j})$ is linear in column k. Consequently, in all cases, $(-1)^{i+j} a_{i\,j} D(A_{i\,j})$ is linear in column k, and thus, $D(A)$ is linear in column k.

Let us now prove that D is alternating. Assume that two adjacent rows of A are equal, say $A^k = A^{k+1}$. First, let $j \neq k$ and $j \neq k+1$. Then, the matrix $A_{i\,j}$ has two identical adjacent columns, and by the induction hypothesis, $D(A_{i\,j}) = 0$. The remaining terms of $D(A)$ are

$$(-1)^{i+k} a_{i\,k} D(A_{i\,k}) + (-1)^{i+k+1} a_{i\,k+1} D(A_{i\,k+1}).$$

However, the two matrices $A_{i\,k}$ and $A_{i\,k+1}$ are equal, since we are assuming that columns k and $k+1$ of A are identical, and since $A_{i\,k}$ is obtained from A by deleting row i and column k, and $A_{i\,k+1}$ is obtained from A by deleting row i and column $k+1$. Similarly, $a_{i\,k} = a_{i\,k+1}$, since columns k and $k+1$ of A are equal. But then,

$$(-1)^{i+k} a_{i\,k} D(A_{i\,k}) + (-1)^{i+k+1} a_{i\,k+1} D(A_{i\,k+1}) = (-1)^{i+k} a_{i\,k} D(A_{i\,k}) - (-1)^{i+k} a_{i\,k} D(A_{i\,k}) = 0.$$

This shows that D is alternating, and completes the proof. $\square$

Lemma 3.5 shows the existence of determinants. We now prove their uniqueness.

Theorem 3.6 *For every $n \geq 1$, for every $D \in \mathcal{D}_n$, for every matrix $A \in \mathrm{M}_n(K)$, we have*

$$D(A) = \sum_{\pi} \epsilon(\pi) a_{\pi(1)\,1} \cdots a_{\pi(n)\,n},$$

where the sum ranges over all permutations π on $\{1, \ldots, n\}$. As a consequence, $\mathcal{D}_n$ consists of a single map for every $n \geq 1$, and this map is given by the above explicit formula.

Proof. Consider the standard basis $(e_1, \ldots, e_n)$ of K^n, where $(e_i)_i = 1$ and $(e_i)_j = 0$, for $j \neq i$. Then, each column A^j of A corresponds to a vector v_j whose coordinates over the basis $(e_1, \ldots, e_n)$ are the components of A^j, that is, we can write

$$v_1 = a_{1\,1} e_1 + \cdots + a_{n\,1} e_n,$$
$$\cdots$$
$$v_n = a_{1\,n} e_1 + \cdots + a_{n\,n} e_n.$$

Since by Lemma 3.5, each D is a multilinear alternating map, by applying Lemma 3.4, we get

$$D(A) = D(v_1, \ldots, v_n) = \Big(\sum_{\pi} \epsilon(\pi) a_{\pi(1)\,1} \cdots a_{\pi(n)\,n}\Big) D(e_1, \ldots, e_n),$$

where the sum ranges over all permutations π on $\{1, \ldots, n\}$. But $D(e_1, \ldots, e_n) = D(I_n)$, and by Lemma 3.5, we have $D(I_n) = 1$. Thus,

$$D(A) = \sum_{\pi} \epsilon(\pi) a_{\pi(1)\,1} \cdots a_{\pi(n)\,n},$$

where the sum ranges over all permutations π on $\{1, \ldots, n\}$. $\square$

We can now prove some properties of determinants.

Corollary 3.7 *For every matrix $A \in \mathrm{M}_n(K)$, we have $D(A) = D(A^\top)$.*

Proof. By Theorem 3.6, we have

$$D(A) = \sum_{\pi} \epsilon(\pi) a_{\pi(1)\,1} \cdots a_{\pi(n)\,n},$$

where the sum ranges over all permutations π on $\{1, \ldots, n\}$. Since a permutation is invertible, every product

$$a_{\pi(1)\,1} \cdots a_{\pi(n)\,n}$$

can be rewritten as

$$a_{1\,\pi^{-1}(1)} \cdots a_{n\,\pi^{-1}(n)},$$

and since $\epsilon(\pi^{-1}) = \epsilon(\pi)$ and the sum is taken over all permutations on $\{1, \ldots, n\}$, we have

$$\sum_{\pi} \epsilon(\pi) a_{\pi(1)\,1} \cdots a_{\pi(n)\,n} = \sum_{\sigma} \epsilon(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)},$$

where π and σ range over all permutations. But it is immediately verified that

$$D(A^\top) = \sum_{\sigma} \epsilon(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)}.$$

$\square$

A useful consequence of Corollary 3.7 is that the determinant of a matrix is also a multilinear alternating map of its rows. This fact, combined with the fact that the determinant of a matrix is a multilinear alternating map of its columns is often useful for finding short-cuts in computing determinants. We illustrate this point on the following example which shows up in polynomial interpolation.

Example 3.2 Consider the so-called *Vandermonde determinant*

$$V(x_1, \ldots, x_n) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{vmatrix}.$$

We claim that

$$V(x_1, \ldots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i),$$

with $V(x_1, \ldots, x_n) = 1$, when $n = 1$. We prove it by induction on $n \geq 1$. The case $n = 1$ is obvious. Assume $n \geq 2$. We proceed as follows: multiply row $n-1$ by x_1 and substract it from row n (the last row), then multiply row $n-2$ by x_1 and subtract it from row $n-1$, etc, multiply row $i-1$ by x_1 and subtract it from row i, until we reach row 1. We obtain the following determinant:

$$V(x_1, \ldots, x_n) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ 0 & x_2 - x_1 & \ldots & x_n - x_1 \\ 0 & x_2(x_2 - x_1) & \ldots & x_n(x_n - x_1) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & x_2^{n-2}(x_2 - x_1) & \ldots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

Now, expanding this determinant according to the first column and using multilinearity, we can factor $(x_i - x_1)$ from the column of index $i-1$ of the matrix obtained by deleting the first row and the first column, and thus

$$V(x_1, \ldots, x_n) = (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) V(x_2, \ldots, x_n),$$

which establishes the induction step.

Lemma 3.4 can be reformulated nicely as follows.

Proposition 3.8 *Let $f \colon E \times \ldots \times E \to F$ be an n-linear alternating map. Let $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ be two families of n vectors, such that*

$$v_1 = a_{1\,1} u_1 + \cdots + a_{1\,n} u_n,$$
$$\ldots$$
$$v_n = a_{n\,1} u_1 + \cdots + a_{n\,n} u_n.$$

Equivalently, letting

$$A = \begin{pmatrix} a_{1\,1} & a_{1\,2} & \ldots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \ldots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n\,1} & a_{n\,2} & \ldots & a_{n\,n} \end{pmatrix}$$

assume that we have

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = A \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Then,

$$f(v_1, \ldots, v_n) = D(A) f(u_1, \ldots, u_n).$$

Proof. The only difference with Lemma 3.4 is that here, we are using $A^\top$ instead of A. Thus, by Lemma 3.4 and Corollary 3.7, we get the desired result. $\square$

As a consequence, we get the very useful property that the determinant of a product of matrices is the product of the determinants of these matrices.

Proposition 3.9 *For any two $n \times n$-matrices A and B, we have $D(AB) = D(A)D(B)$.*

Proof. We use Proposition 3.8 as follows: let $(e_1, \ldots, e_n)$ be the standard basis of K^n, and let

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = AB \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}.$$

Then, we get

$$D(w_1, \ldots, w_n) = D(AB) D(e_1, \ldots, e_n) = D(AB),$$

since $D(e_1, \ldots, e_n) = 1$. Now, letting

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = B \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix},$$

we get

$$D(v_1, \ldots, v_n) = D(B),$$

and since

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = A \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

we get

$$D(w_1, \ldots, w_n) = D(A) D(v_1, \ldots, v_n) = D(A) D(B).$$

□

It should be noted that all the results of this section, up to now, also holds when K is a commutative ring, and not necessarily a field. We can now characterize when an $n \times n$-matrix A is invertible in terms of its determinant $D(A)$.

3.4 Inverse Matrices and Determinants

In the next two sections, K is a commutative ring and when needed, a field.

Definition 3.6 Let K be a commutative ring. Given a matrix $A \in \mathrm{M}_n(K)$, let $\widetilde{A} = (b_{i\,j})$ be the matrix defined such that

$$b_{i\,j} = (-1)^{i+j} D(A_{j\,i}),$$

the cofactor of $a_{j\,i}$. The matrix $A_{j\,i}$ is called a *minor* of the matrix A.

Note the reversal of the indices in

$$b_{i\,j} = (-1)^{i+j} D(A_{j\,i}).$$

Thus, $\widetilde{A}$ is the transpose of the matrix of cofactors of elements of A.

We have the following proposition.

Proposition 3.10 *Let K be a commutative ring. For every matrix $A \in \mathrm{M}_n(K)$, we have*

$$A\widetilde{A} = \widetilde{A}A = \det(A) I_n.$$

As a consequence, A is invertible iff $\det(A)$ is invertible, and if so, $A^{-1} = (\det(A))^{-1}\widetilde{A}$.

Proof. If $\widetilde{A} = (b_{i\,j})$ and $A\widetilde{A} = (c_{i\,j})$, we know that the entry $c_{i\,j}$ in row i and column j of $A\widetilde{A}$ is

$$c_{i\,j} = a_{i\,1} b_{1\,j} + \cdots + a_{i\,k} b_{k\,j} + \cdots + a_{i\,n} b_{n\,j},$$

which is equal to

$$a_{i\,1}(-1)^{j+1} D(A_{j\,1}) + \cdots + a_{i\,n}(-1)^{j+n} D(A_{j\,n}).$$

If $j = i$, then we recognize the expression of the expansion of $\det(A)$ according to the i-th row:

$$c_{i\,i} = \det(A) = a_{i\,1}(-1)^{i+1} D(A_{i\,1}) + \cdots + a_{i\,n}(-1)^{i+n} D(A_{i\,n}).$$

If $j \neq i$, we can form the matrix A' by replacing the j-th row of A by the i-th row of A. Now, the matrix $A_{j\,k}$ obtained by deleting row j and column k from A is equal to the matrix

$A'_{j\,k}$ obtained by deleting row j and column k from A', since A and A' only differ by the j-th row. Thus,

$$D(A_{j\,k}) = D(A'_{j\,k}),$$

and we have

$$c_{i\,j} = a_{i\,1}(-1)^{j+1}D(A'_{j\,1}) + \cdots + a_{i\,n}(-1)^{j+n}D(A'_{j\,n}).$$

However, this is the expansion of $D(A')$ according to the j-th row, since the j-th row of A' is equal to the i-th row of A, and since A' has two identical rows i and j, because D is an alternating map of the rows (see an earlier remark), we have $D(A') = 0$. Thus, we have shown that $c_{i\,i} = \det(A)$, and $c_{i\,j} = 0$, when $j \neq i$, and so

$$A\widetilde{A} = \det(A)I_n.$$

It is also obvious from the definition of $\widetilde{A}$, that

$$\widetilde{A}^{\top} = \widetilde{A^{\top}}.$$

Then, applying the first part of the argument to $A^{\top}$, we have

$$A^{\top}\widetilde{A^{\top}} = \det(A^{\top})I_n,$$

and since, $\det(A^{\top}) = \det(A)$, $\widetilde{A}^{\top} = \widetilde{A^{\top}}$, and $(\widetilde{A}A)^{\top} = A^{\top}\widetilde{A}^{\top}$, we get

$$\det(A)I_n = A^{\top}\widetilde{A^{\top}} = A^{\top}\widetilde{A}^{\top} = (\widetilde{A}A)^{\top},$$

that is,

$$(\widetilde{A}A)^{\top} = \det(A)I_n,$$

which yields

$$\widetilde{A}A = \det(A)I_n,$$

since $I_n^{\top} = I_n$. This proves that

$$A\widetilde{A} = \widetilde{A}A = \det(A)I_n.$$

As a consequence, if $\det(A)$ is invertible, we have $A^{-1} = (\det(A))^{-1}\widetilde{A}$. Conversely, if A is invertible, from $AA^{-1} = I_n$, by Proposition 3.9, we have $\det(A)\det(A^{-1}) = 1$, and $\det(A)$ is invertible. □

When K is a field, an element $a \in K$ is invertible iff $a \neq 0$. In this case, the second part of the proposition can be stated as A is invertible iff $\det(A) \neq 0$. Note in passing that this method of computing the inverse of a matrix is usually not practical.

We now consider some applications of determinants to linear independence and to solving systems of linear equations. Although these results hold for matrices over a commutative ring, their proofs require more sophisticated methods than when K is a field. Therefore, we assume again that K is a field.

Let A be an $n \times n$-matrix, X a column vectors of variables, and B another column vector, and let $A^1, \ldots, A^n$ denote the columns of A. Observe that the system of equation $AX = B$,

$$\begin{pmatrix} a_{1\,1} & a_{1\,2} & \ldots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \ldots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n\,1} & a_{n\,2} & \ldots & a_{n\,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

is equivalent to

$$x_1 A^1 + \cdots + x_j A^j + \cdots + x_n A^n = B,$$

since the equation corresponding to the i-th row is in both cases

$$a_{i\,1} x_1 + \cdots + a_{i\,j} x_j + \cdots + a_{i\,n} x_n = b_i.$$

First, we characterize linear independence of the column vectors of a matrix A in terms of its determinant.

Proposition 3.11 *Given an $n \times n$-matrix A over a field K, the columns $A^1, \ldots, A^n$ of A are linearly dependent iff $D(A) = D(A^1, \ldots, A^n) = 0$. Equivalently, A has rank n iff $D(A) \neq 0$.*

Proof. First, assume that the columns $A^1, \ldots, A^n$ of A are linearly dependent. Then, there are $x_1, \ldots, x_n \in K$, such that

$$x_1 A^1 + \cdots + x_j A^j + \cdots + x_n A^n = 0,$$

where $x_j \neq 0$ for some j. If we compute

$$D(A^1, \ldots, x_1 A^1 + \cdots + x_j A^j + \cdots + x_n A^n, \ldots, A^n) = D(A^1, \ldots, 0, \ldots, A^n) = 0,$$

where 0 occurs in the j-th position, by multilinearity, all terms containing two identical columns A^k for $k \neq j$ vanish, and we get

$$x_j D(A^1, \ldots, A^n) = 0$$

for every j, $1 \leq j \leq n$. Since $x_j \neq 0$ for some j, and K is a field, we must have $D(A^1, \ldots, A^n) = 0$.

Conversely, we show that if the columns $A^1, \ldots, A^n$ of A are linearly independent, then $D(A^1, \ldots, A^n) \neq 0$. If the columns $A^1, \ldots, A^n$ of A are linearly independent, then they form a basis of K^n, and we can express the standard basis $(e_1, \ldots, e_n)$ of K^n in terms of $A^1, \ldots, A^n$. Thus, we have

$$\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} b_{1\,1} & b_{1\,2} & \ldots & b_{1\,n} \\ b_{2\,1} & b_{2\,2} & \ldots & b_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n\,1} & b_{n\,2} & \ldots & b_{n\,n} \end{pmatrix} \begin{pmatrix} A^1 \\ A^2 \\ \vdots \\ A^n \end{pmatrix},$$

for some matrix $B = (b_{i\,j})$, and by Proposition 3.8, we get

$$D(e_1, \ldots, e_n) = D(B)D(A^1, \ldots, A^n),$$

and since $D(e_1, \ldots, e_n) = 1$, this implies that $D(A^1, \ldots, A^n) \neq 0$ (and $D(B) \neq 0$). For the second assertion, recall that the rank of a matrix is equal to the maximum number of linearly independent columns, and the conclusion is clear. $\square$

3.5 Systems of Linear Equations and Determinant

We now characterize when a system of linear equations of the form $AX = B$ has a unique solution.

Proposition 3.12 *Given an $n \times n$-matrix A over a field K, the following properties hold:*

(1) For every column vector B, there is a unique column vector X such that $AX = B$ iff the only solution to $AX = 0$ is the trivial vector $X = 0$, iff $D(A) \neq 0$.

(2) If $D(A) \neq 0$, the unique solution of $AX = B$ is given by the expressions

$$X_j = \frac{D(A^1, \ldots, A^{j-1}, B, A^{j+1}, \ldots, A^n)}{D(A^1, \ldots, A^{j-1}, A^j, A^{j+1}, \ldots, A^n)},$$

known as Cramer's rules.

(3) The system of linear equations $AX = 0$ has a nonzero solution iff $D(A) = 0$.

Proof. Assume that $AX = B$ has a single solution X_0, and assume that $AY = 0$ with $Y \neq 0$. Then,

$$A(X_0 + Y) = AX_0 + AY = AX_0 + 0 = B,$$

and $X_0 + Y \neq X_0$ is another solution of $AX = B$, contadicting the hypothesis that $AX = B$ has a single solution X_0. Thus, $AX = 0$ only has the trivial solution. Now, assume that $AX = 0$ only has the trivial solution. This means that the columns $A^1, \ldots, A^n$ of A are linearly independent, and by Proposition 3.11, we have $D(A) \neq 0$. Finally, if $D(A) \neq 0$, by Proposition 3.10, this means that A is invertible, and then, for every B, $AX = B$ is equivalent to $X = A^{-1}B$, which shows that $AX = B$ has a single solution.

(2) Assume that $AX = B$. If we compute

$$D(A^1, \ldots, x_1 A^1 + \cdots + x_j A^j + \cdots + x_n A^n, \ldots, A^n) = D(A^1, \ldots, B, \ldots, A^n),$$

where B occurs in the j-th position, by multilinearity, all terms containing two identical columns A^k for $k \neq j$ vanish, and we get

$$x_j D(A^1, \ldots, A^n) = D(A^1, \ldots, A^{j-1}, B, A^{j+1}, \ldots, A^n),$$

for every j, $1 \leq j \leq n$. Since we assumed that $D(A) = D(A^1, \ldots, A^n) \neq 0$, we get the desired expression.

(3) Note that $AX = 0$ has a nonzero solution iff $A^1, \ldots, A^n$ are linearly dependent (as observed in the proof of Proposition 3.11), which, by Proposition 3.11, is equivalent to $D(A) = 0$. □

As pleasing as Cramer's rules are, it is usually impractical to solve systems of linear equations using the above expressions.

3.6 Determinant of a Linear Map

We close this chapter with the notion of determinant of a linear map $f: E \to E$.

Given a vector space E of finite dimension n, given a basis $(u_1, \ldots, u_n)$ of E, for every linear map $f: E \to E$, if $M(f)$ is the matrix of f w.r.t. the basis $(u_1, \ldots, u_n)$, we can define $\det(f) = \det(M(f))$. If $(v_1, \ldots, v_n)$ is any other basis of E, and if P is the change of basis matrix, by Corollary 2.15, the matrix of f with respect to the basis $(v_1, \ldots, v_n)$ is $P^{-1}M(f)P$. Now, by proposition 3.9, we have

$$\det(P^{-1}M(f)P) = \det(P^{-1})\det(M(f))\det(P) = \det(P^{-1})\det(P)\det(M(f)) = \det(M(f)).$$

Thus, $\det(f)$ is indeed independent of the basis of E.

Definition 3.7 Given a vector space E of finite dimension, for any linear map $f: E \to E$, we define the *determinant* $\det(f)$ *of* f as the determinant $\det(M(f))$ of the matrix of f in any basis (since, from the discussion just before this definition, this determinant does not depend on the basis).

Then, we have the following proposition.

Proposition 3.13 *Given any vector space E of finite dimension n, a linear map $f: E \to E$ is invertible iff* $\det(f) \neq 0$.

Proof. The linear map $f: E \to E$ is invertible iff its matrix $M(f)$ in any basis is invertible (by Proposition 2.12), iff $\det(M(f)) \neq 0$, by Proposition 3.10.

Given a vector space of finite dimension n, it is easily seen that the set of bijective linear maps $f: E \to E$ such that $\det(f) = 1$ is a group under composition. This group is a subgroup of the general linear group $\mathbf{GL}(E)$. It is called the *special linear group (of E)*, and it is denoted by $\mathbf{SL}(E)$, or when $E = K^n$, by $\mathbf{SL}(n, K)$, or even by $\mathbf{SL}(n)$.

3.7 Gausian Elimination and LU Decomposition

3.8 Futher Readings

Other sources include:

Linear Algebra and its Applications, by Strang [31], *Algebra*, by Lang [18], *Algebra*, by Michael Artin [1], *Algebra*, by Mac Lane and Birkhoff [20], *Algèbre*, by Bourbaki [7, 8], *Algebra 1*, by Van Der Waerden [33], *Algèbre linéaire et géométrie classique*, by Bertin [6], *Géométrie 1 and 2*, by Berger [3, 4], *Géométries affines, projectives, and euclidiennes*, by Tisseron [32], *Algèbre Linéaire et Géométrie Elémentaire* by Dieudonné [11], for linear algebra, polynomials, and geometry,

Cours de Calcul Différentiel, by Henri Cartan [9], the famous Analysis Courses *Analyse I-IV*, by Schwartz [26, 27, 28, 29], *Calcul différentiel*, by Avez [2], *Analysis on Manifolds*, by Munkres [22], highly recommended as a very clear and readable text, for basic topology and differential calculus, and

Géométrie différentielle: variétés, courbes et surfaces, by Berger and Gostiaux, [5], *Differential Geometry of Curves and Surfaces*, by do Carmo [12], a beautiful classic on the subject, for the geometry of curves and surfaces.

Another useful and rather complete reference is the text *Finite-Dimensional Spaces*, by Walter Noll [23]. But beware, the notation and terminology is a bit strange!

We now turn to polynomials.